Holistic Cost Analysis of Running a Computing Center

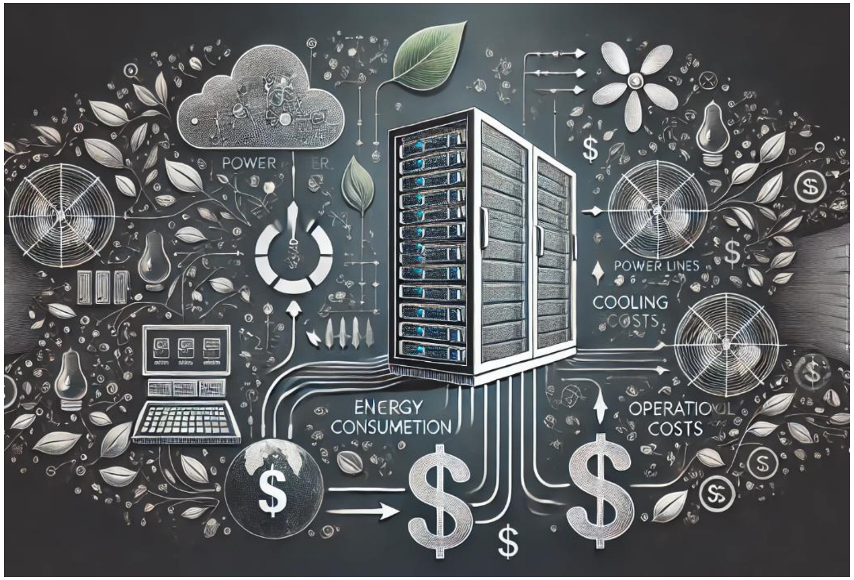
Christoph Paus, Maxim Goncharov, Zhangqier Wang

WLCG Workshop

Dec. 11th 2024

Introduction

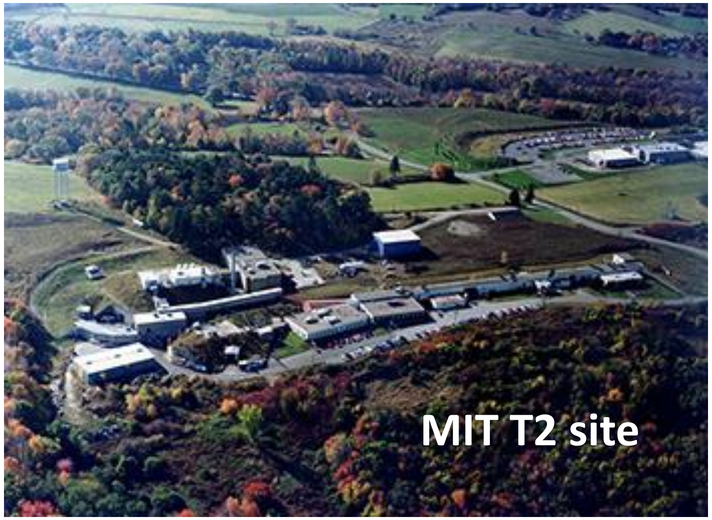
- The increasing volume of data (ex. HL-LHC), new physics exploration, and AI applications are driving a surge in computing resource requirements
- Running a computing center involves multifaceted costs; make a holistic cost analysis essential for long-term operation.
- Beyond performance metrics, more factors to consider
  - **Power consumption** including cooling, a key contributor to operational expenses
  - Physical space usage
- Computing in general faces a major re-design to align with energy efficiency and sustainability goals.



Worldwide LHC Computing Grid

MIT Tier 2 Center

- The MIT Tier-2 Center is a high-performance computing facility dedicated to processing, storing, and analyzing data for the CMS, LHCb, and other experiments
  - Comprises approximately **700 machines**, providing **25k CPU cores** and **16.5 PB of storage**.
  - Compute/Storage mix model
- Worker nodes are also used as storage devices
  - Re-design to have dedicated compute and storage servers
- Cost evaluation determine hardware retirement policy
  - Prepare MIT T2 for HL-LHC (data x10)
  - Improve energy efficiency
  - Spend less to provide same amount of computation



WLCG
Worldwide LHC Computing Grid

Machine Categories

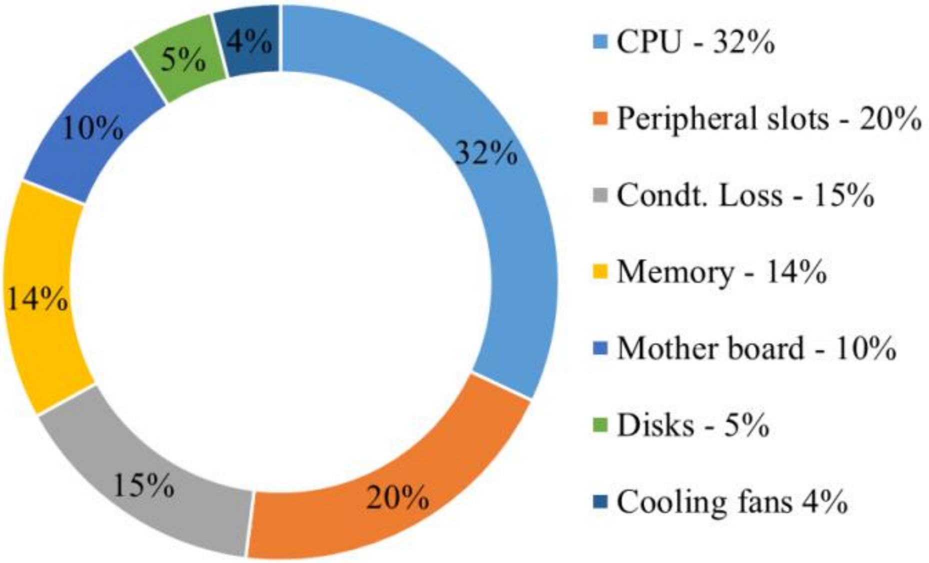
- CPU models categorized into 8 types
  - power consumption, and cpu, memory usage are checked
  - average year represents the age of the machine

CPU model	Avg year	Process node	HS06	Cores	HS06/core
Intel(R)_Xeon(R)_E5310-5410	2008	65 nm	69	8	8.6
Intel(R)_Xeon(R)_X5647	2017	32 nm	155	16	9.7
Intel(R)_Xeon(R)_E5520-5620	2018	45 nm	120-140	16	8.1
Intel(R)_Xeon(R)_E5-series	2018	14/22 nm	169-449	8-40	11.1
Intel(R)_Xeon(R)_Silver	2019	14 nm	530-706	48-64	11.0
Intel(R)_Xeon(R)_Gold	2021	10 nm	904	64	14.1
AMD_EPYC_9754_128-Core_Processor	2023	5 nm	7450	512	14.6

Typical Power Consumption

- With higher computing usage nowadays, the cost of power consumption plays higher part.
- The power usage consists of multiple parts, with CPU usage being the major contribution.

- CPU usage fraction differs between different CPU models
- At MIT T2, the fraction of CPU + memory usage is 40%-60%.
- The newer CPU model has more cores per machine and has more fraction of power used for CPU usage.
- GPU is more power hungry



ACCESS.2021.3125092

Worldwide LHC Computing Grid

Power Cross-check

- The power consumption is monitored using “ipmitool” and “omreport”.
- The current is measured on two servers and compared to the current from the monitoring
  - Measured using clamp meter and AC splitter
  - Load CPUs using linux stress command “stress --cpu N --timeout 100”
  - Load CPU with CMS actually process using 16/32 cores.

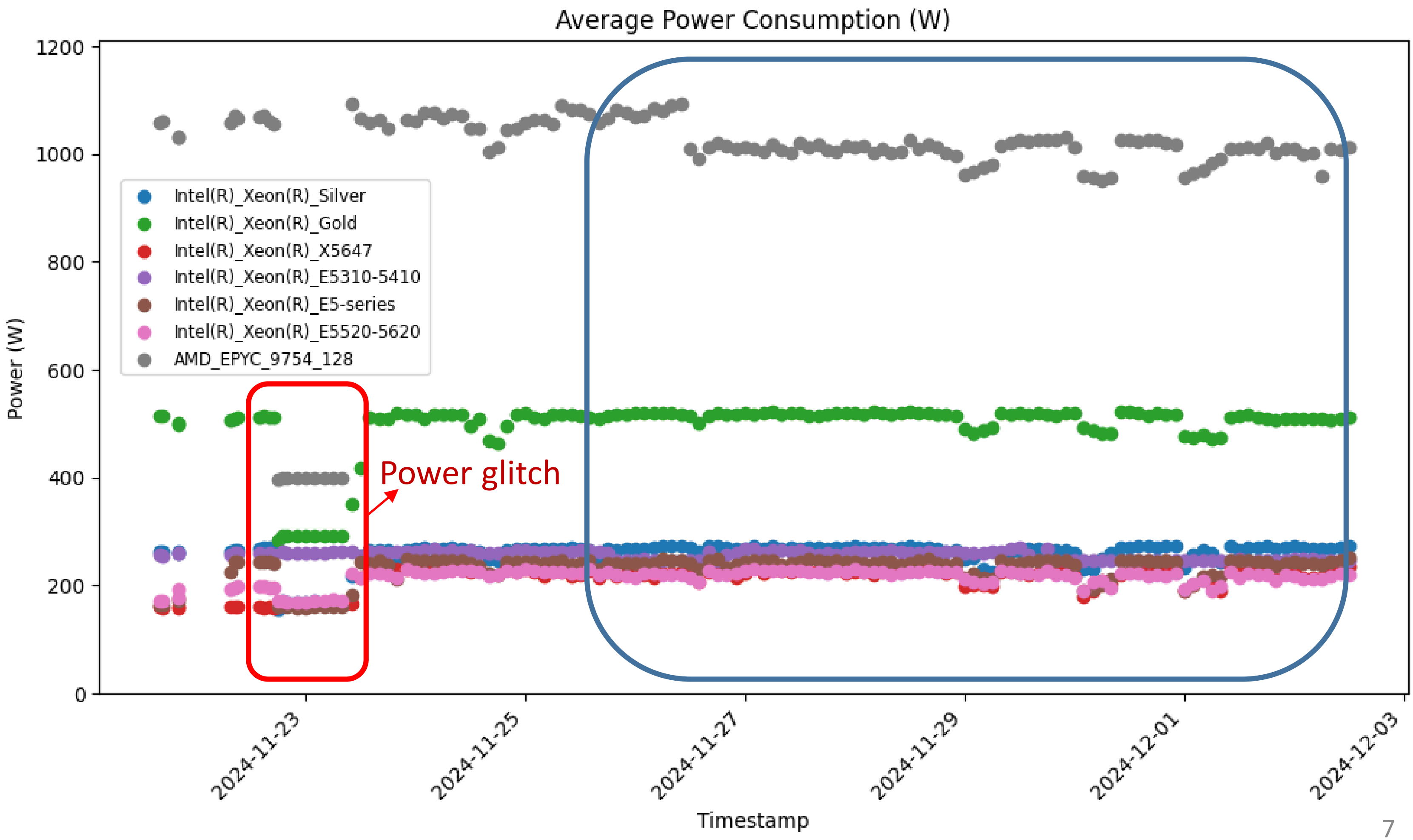
Current (A)		Base	16 cpu	32 cpu	48 cpu	64 cpu	CMS 16 cores	CMS 32 cores
Server1	Meter	2.80	3.72	4.44	4.63	4.81	4.07	5.00
	Monitor	2.8	3.8	4.4	4.6	4.8	4.0	5.0
Server2	Meter	2.49	3.44	4.13	4.37	4.49	3.76	4.64
	monitor	2.4	3.4	4.0	4.2	4.4	3.7	4.6

vs

Consistent current reading from monitoring and measurements

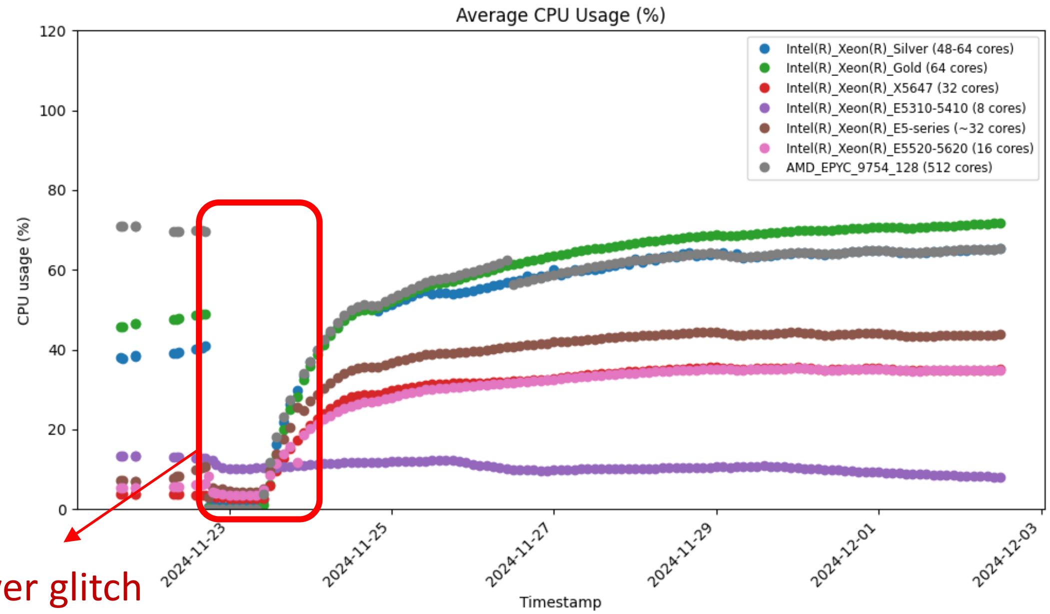
MIT Tier 2 Power Usage

- Power consumption is relatively stable for Tier 2 operation.
  - Cost analysis based on data from the "plateau" region Nov.26th -Dec 2nd

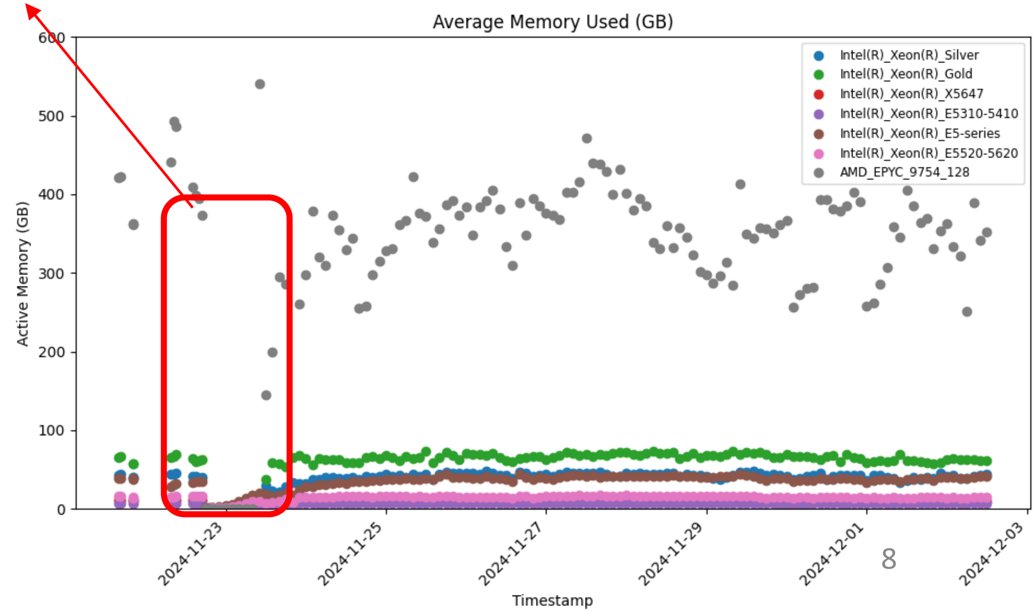


MIT Tier 2 CPU usage

- Memory usage extracted from active memory via 'vmstats –s'
- CPU usage and memory usage is highly correlated
  - average 1.2 GB/core
- Power consumption and CPU usage is highly correlated.
- Estimate the computing resource and its connection to the power consumption.
  - Evaluated by Power/HS06



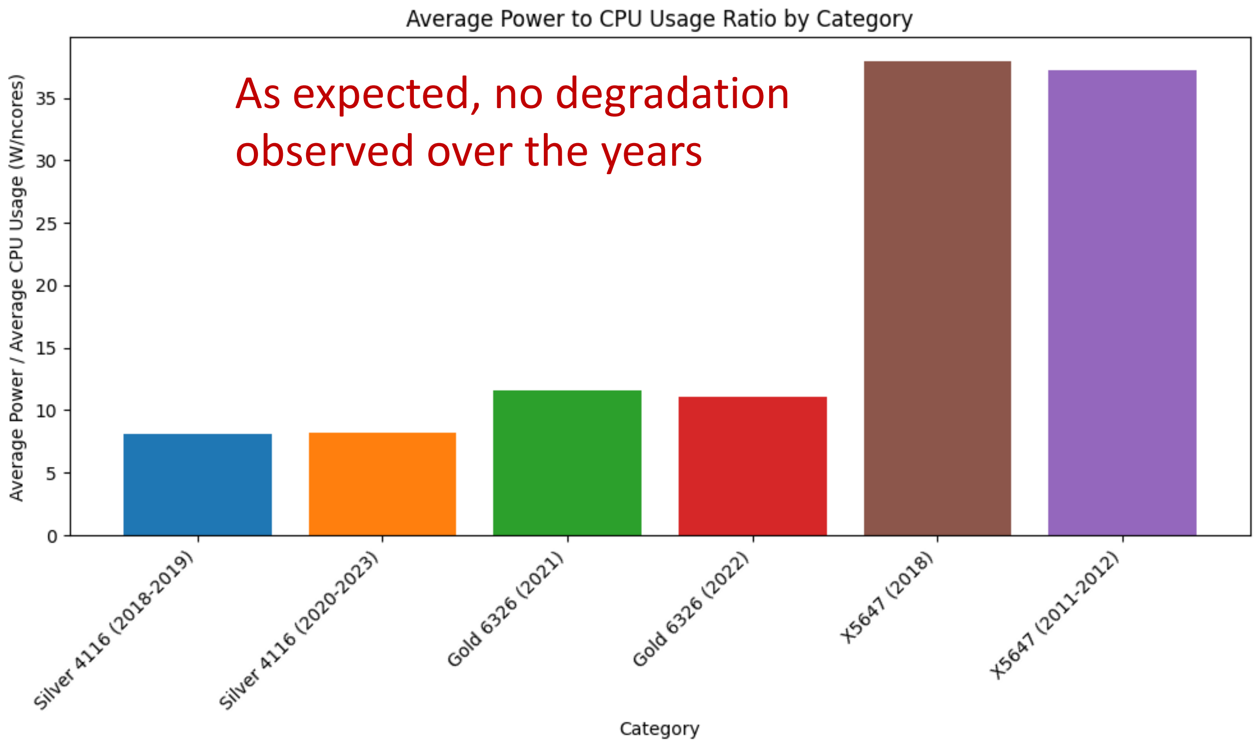
Power glitch



WLCG
Worldwide LHC Computing Grid

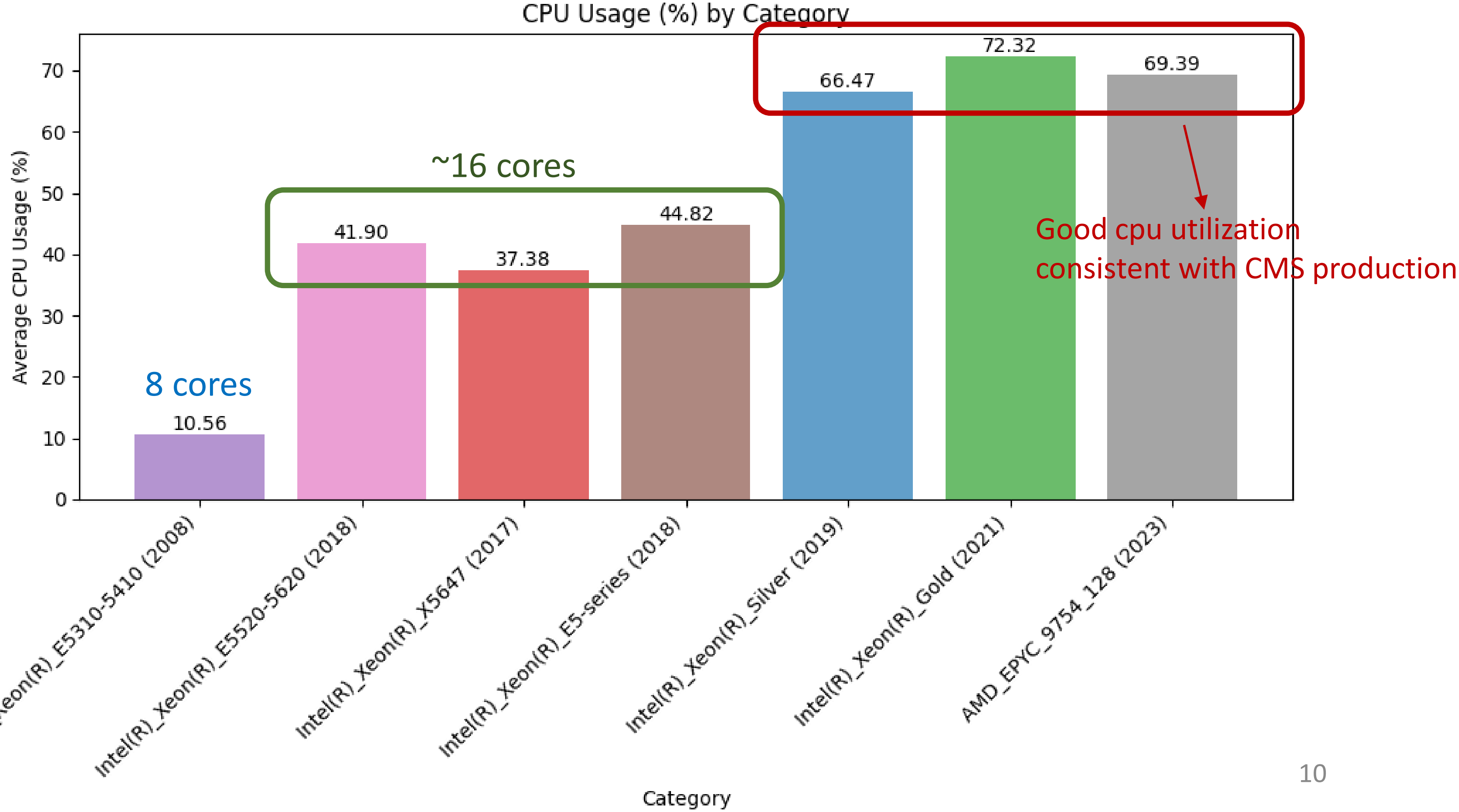
Machine Age

- Machines with the same CPU model may be produced in different years and have been running longer.
  - Aging effect on power consumption is checked
- Check the power usage on 3 CPU models
  - Intel(R)_Xeon(R)_Silver_4116_2.10GHz: 104 machines (2018-2019), 2 machines (2021-2023)
  - Intel(R)_Xeon(R)_Gold_6326_2.90GHz: 36 machines (2021), 22 machines (2022)
  - Intel(R)_Xeon(R)_X5647_2.93GHz: 2 machines (2011-2012), 63 machines (2017)



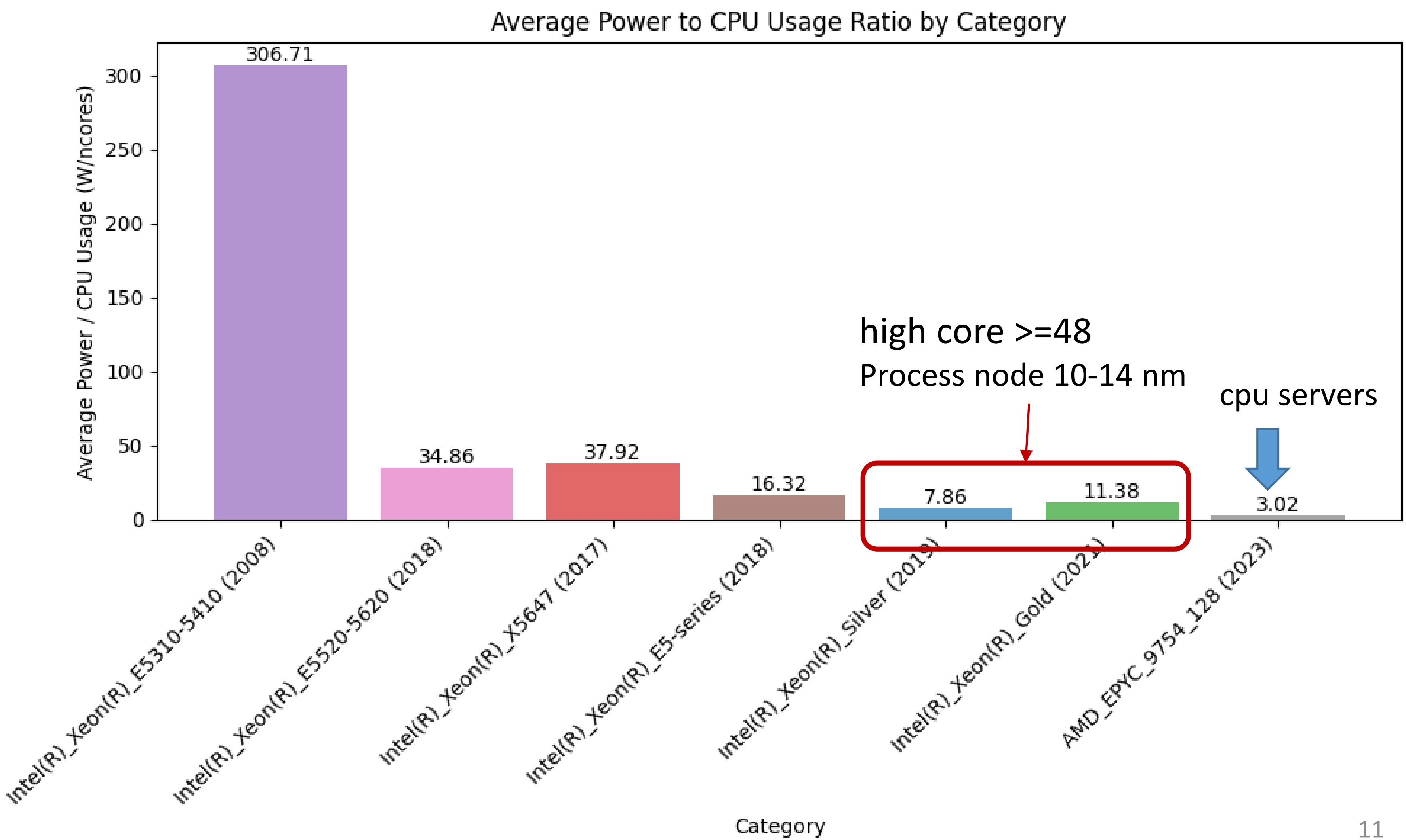
CPU Usage Comparison

- Old machines tend to be less used, due to job mismatches
  - Intel(R)_Xeon(R)_E5310-5410 have only 8 cores per machine, no longer suitable for modern computation needs
  - Production pilots using 8 cores, CPU usage is very low for low core machines



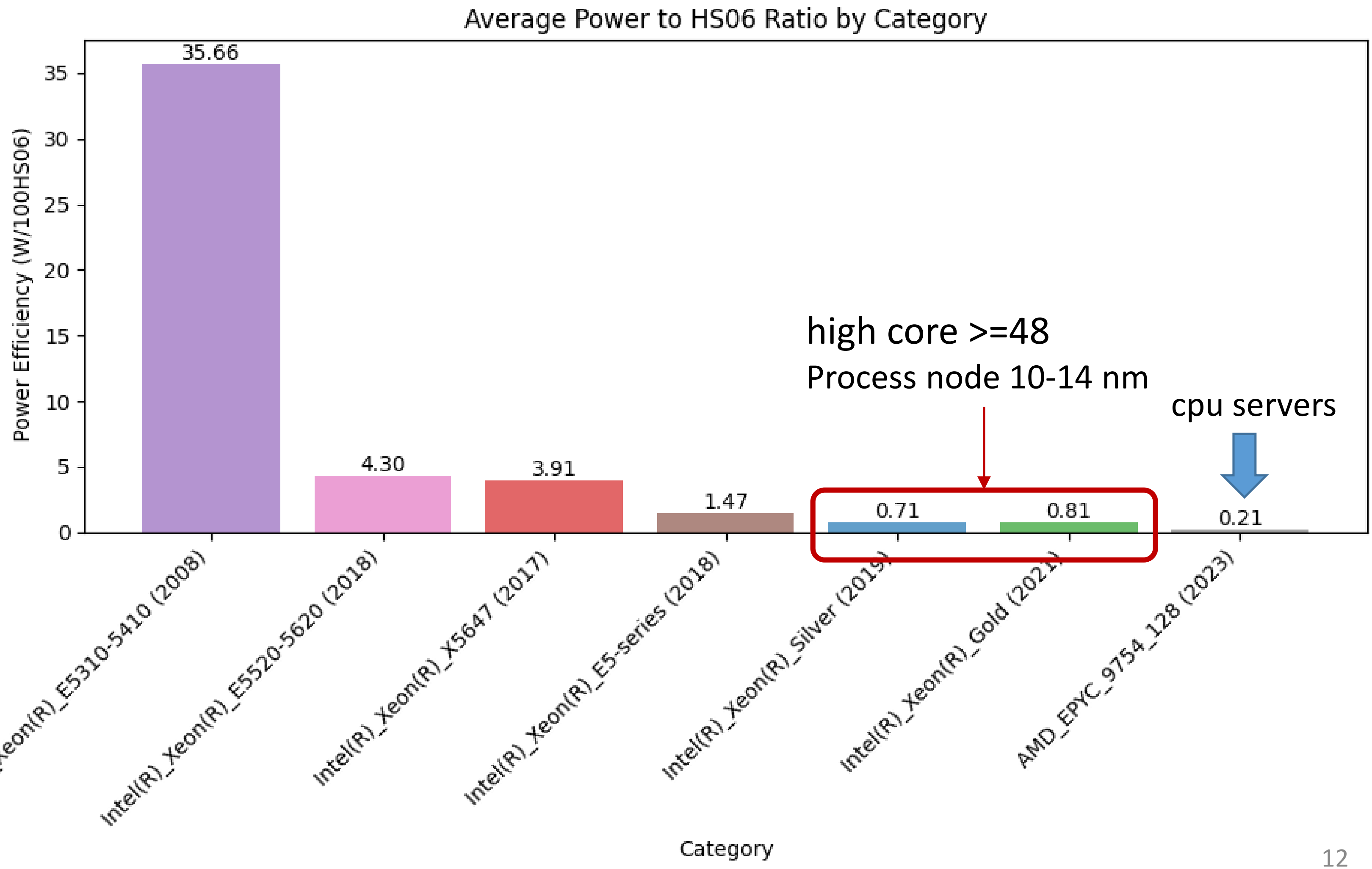
Power/Core Comparison

- The average power consumption for delivering 1 computing core



Power/HS06 Comparison

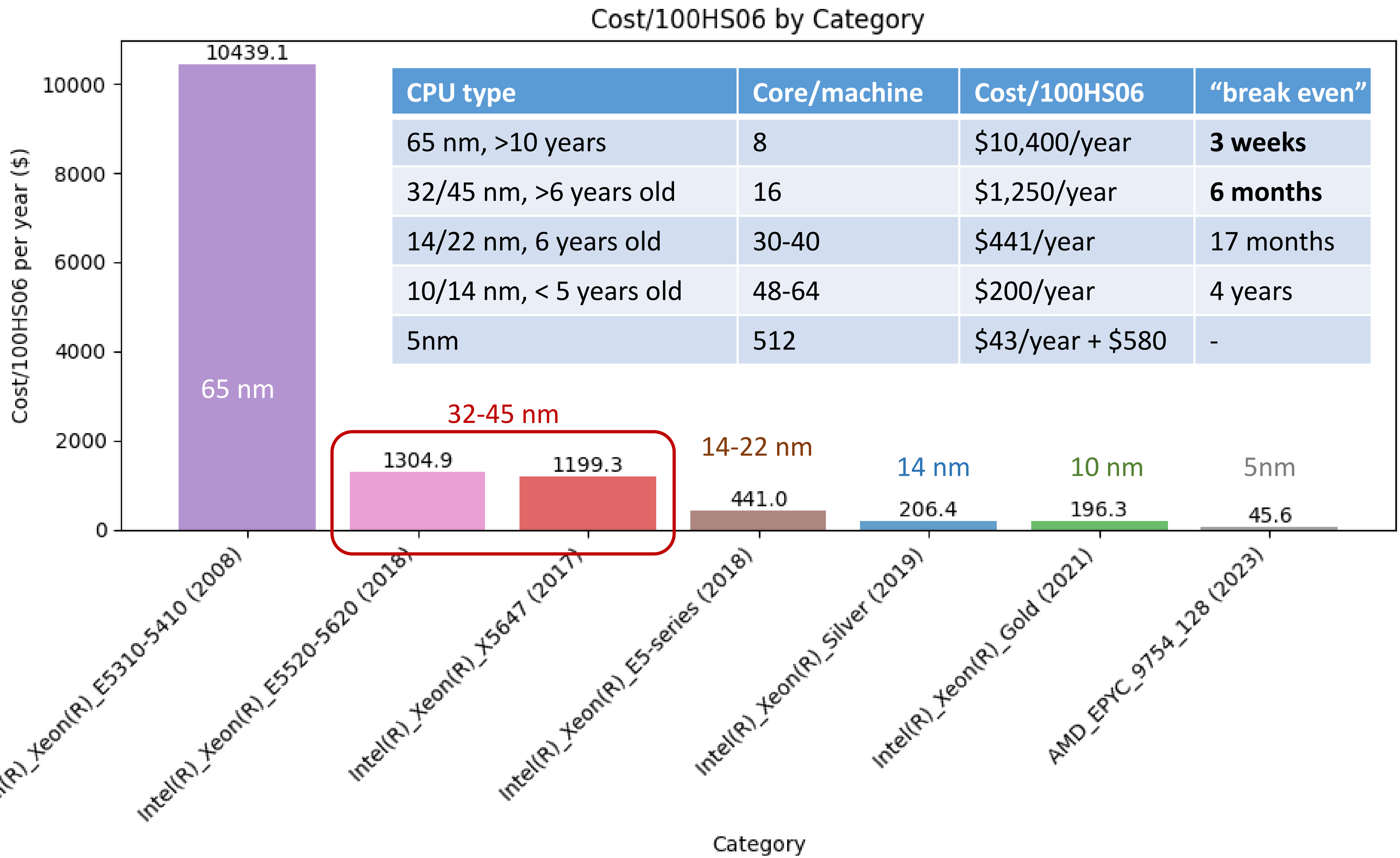
- The average power consumption for delivering 100 HS06 of compute



Example Cost Analysis

- For each type of machines, the cost includes power, space, and cooling.
  - Power price at T2: 14 – 18 cents / kWh
  - Power usage effectiveness (PUE) is 1.4 as a typical example (not very good)
  - Space usage: >$5000 / (40 unit rack) every year
  - Yearly cost = PUE * power * $0.16/kWh * year + $5000 * (rack space)
- **Cost of providing 100 HS06 computation**
  - **Replace with new CPU server** AMD_EPYC_9754 (**5nm**)
    - Cost = $580 (purchase) + $43/year
    - 100 HS06 is provided by 1.9% of a single server
    - Among $43 per year, $2.5 comes from rack usage, $41 comes from power bill.
  - Intel(R)_Xeon(R)_E5310-5410 (**65nm, >10 years old**)
    - Cost = $10,400/year
    - If replaced, after 3 weeks it will break even
  - Similar estimations for other CPU models

Cost Summary



CPU type	Core/machine	Cost/100HS06	"break even"
65 nm, >10 years	8	$10,400/year	**3 weeks**
32/45 nm, >6 years old	16	$1,250/year	**6 months**
14/22 nm, 6 years old	30-40	$441/year	17 months
10/14 nm, < 5 years old	48-64	$200/year	4 years
5nm	512	$43/year + $580	-

break even *: make a positive reform of investment

Conclusion

- Holistic Cost Analysis is presented for running a computing center for HEP
  - **Old hardware** should be replaced to reduce long-term operational costs
  - Cost saving can be substantial
- Hardware replacement rule of thumb
  - Hardware older than 10 years need replacing immediately
    - **Even turning off helps save the cost**
  - Low core machines (< 32 cores) older than 6 years worth replacing
    - Process node > 30 nm
    - Saving cost after 1-2 years of operation
  - Machines older than 5 years could be considered
    - Depending on power prices and PUE at the site
    - Generally saving cost after 2-4 years of operation

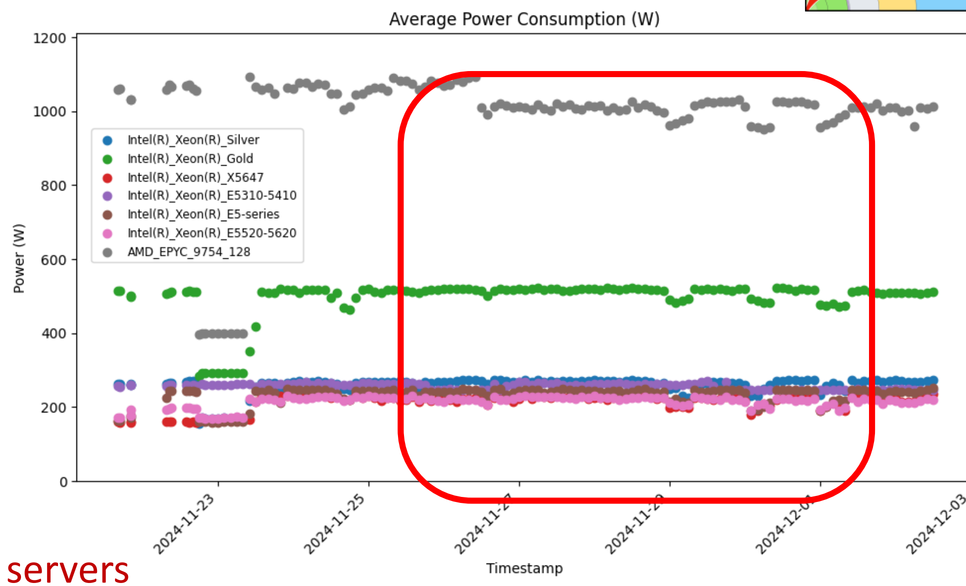
Worldwide LHC Computing Grid

The End

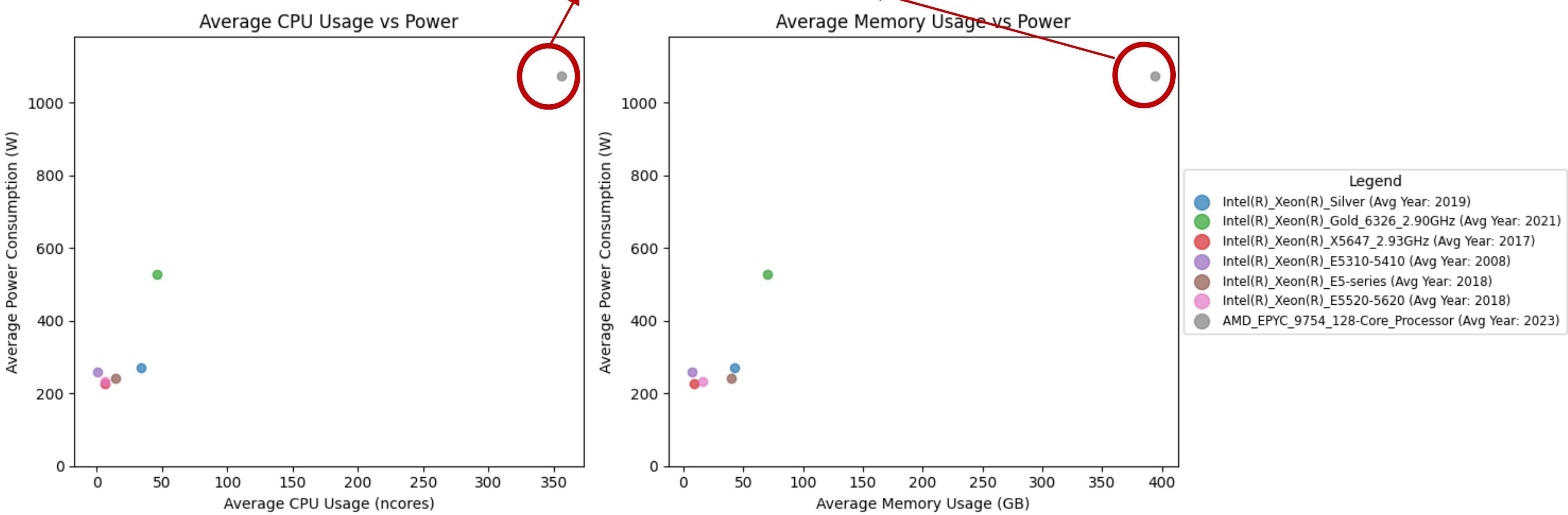
Back up

Scatter Plots

- Average usage vs power extracted from the "plateau" region
  - Nov.26th -Dec 2nd
- Newer CPU model tend to have larger power/cpu usage

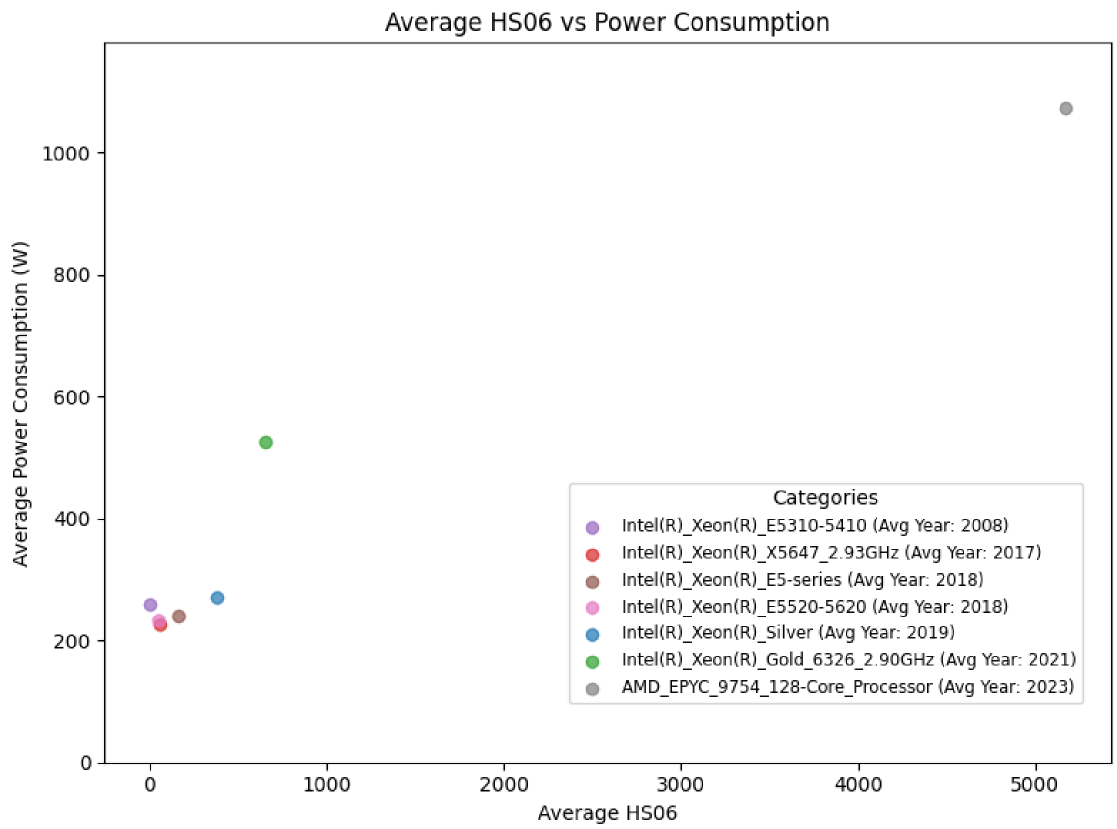


Dedicated cpu servers



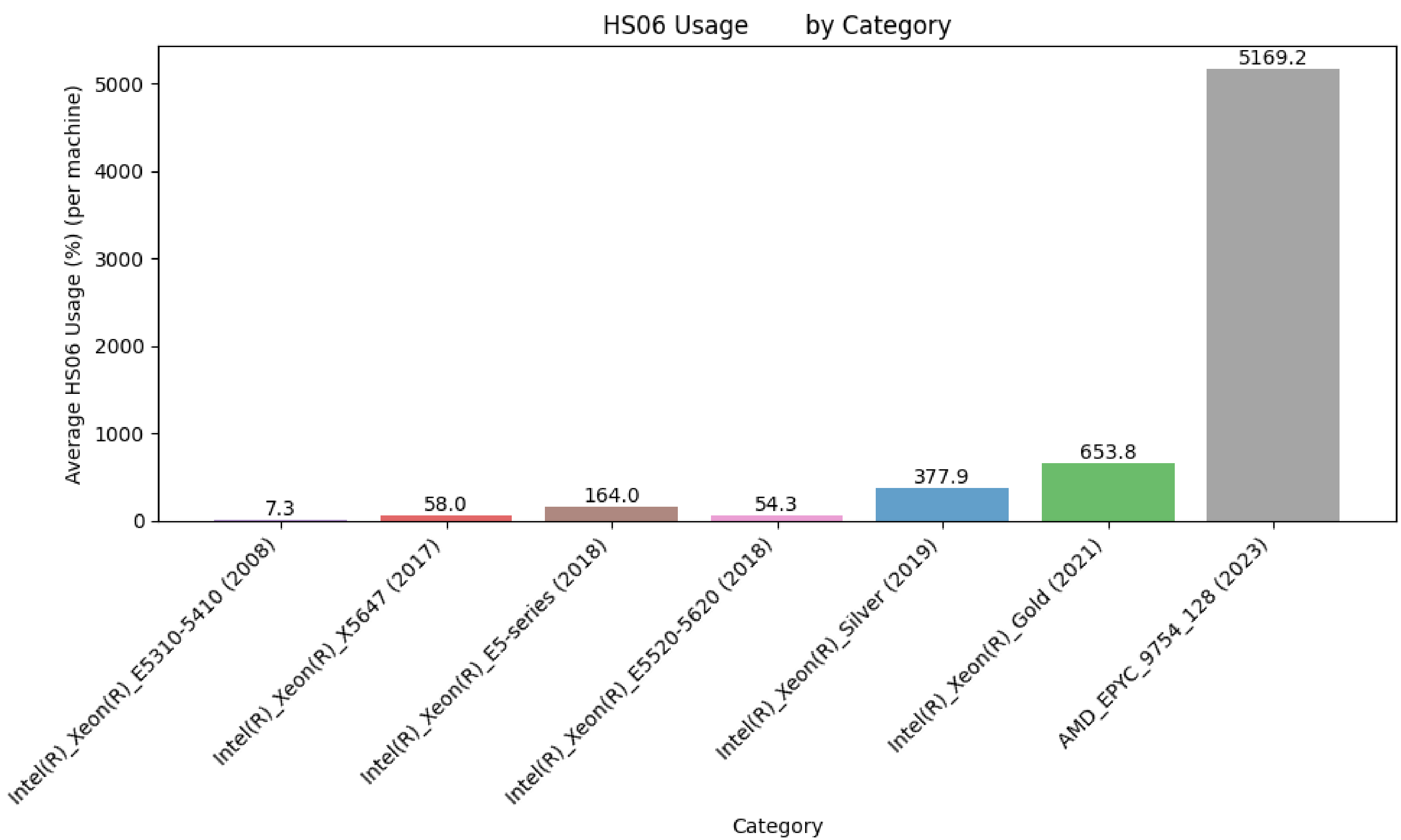
Scatter

- The average HS06 per machine



HS06 Comparison

- The average HS06 per machine



Cost Summary

- **Cost of providing 100 HS06 computation**
  - Intel(R)_Xeon(R)_X5647, E5520-5620 (32/45 nm, 2017-2018)
    - Cost = 1250$/year
    - If replaced, saving cost after 6 months running
  - Intel(R)_Xeon(R)_E5-series (14/22 nm, 2018)
    - Cost = 441$/year
    - If replaced, saving cost in less than 2 years running
  - Intel(R)_Xeon(R)_Silver, Gold (10/14 nm, 2019-2021)
    - Cost =200$/year
    - If replaced, saving cost after 3.7 years running

CPU type	Core/machine	Cost/100HS06	time for cost saving
65 nm, >10 years	8	10,400$/year	**3 weeks**
32/45 nm, >6 years old	16	1,250$/year	**6 months**
14/22 nm, 6 years old	30-40	441$/year	17 months
10/14 nm, < 5 years old	48-64	200$/year	4 years
5nm	512	43$/year + 580$	-